SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022

Multi-Stage Stochastic Linear Programs for Portfolio Optimization

by
George B. Dantzig and Gerd Infanger

TECHNICAL REPORT SOL 91-11

September 1991

Research and reproduction of this report were partially supported by the Office of Naval Research Grant N00014-89-J-1659; the National Science Foundation Grants ECS-8906260, DMS-8913089; and the Electric Power Research Institute Contracts RP 8010-09 and CSA-4O05335, and the Austrian Science Foundation,"Fonds zur Förderung der wissenschaftlichen Forschung," Grant J0323-PHY.

This report is being presented and published at the Proceedings of the Annual Symposium of RAMP (Research Association of Mathematical Programming) in Tokyo, November 1991.

Multi-Stage Stochastic Linear Programs for Portfolio Optimization*

George B. Dantzig† and Gerd Infanger†

Abstract

The paper demonstrates how multi-period portfolio optimization problems can be efficiently solved as multi-stage stochastic linear programs. A scheme based on a blending of classical Benders decomposition techniques and a special technique, called importance sampling, is used to solve this general class of multi-stage stochastic linear programs. We discuss the case where stochastic parameters are dependent within a period as well as between periods. Initial computational results are presented.

*Research and reproduction of this report were partially supported by the Office of Naval Research Contract N00014-89-J-1659; the National Science Foundation Grants ECS-8906260, DMS-8913089, the Electric Power Research Institute Contract RP 8010-09, CSA-4O05335, and the Austrian Science Foundation, "Fonds zur Förderung der wissenschaftlichen Forschung," Grant J0323-Phy. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do NOT necessarily reflect the views of the above sponsors.

†Department of Operations Research, Stanford University, Stanford, CA 94305-4022, U.S.A.

1. Introduction

Methods of Operations Research, especially Mathematical Programming methods, are receiving broader acceptance in the financial industry. The increasing complexities and inherent uncertainties in financial markets have lead to the need of mathematical models supporting the decision making process. This paper addresses the portfolio selection problem. Since Markowitz (1959) [20], several models have been developed that allow one to determine portfolios with the highest expected returns for a given level of risk. His model (and certain closely related ones) require the solution of a quadratic program. Other approaches model the stochastic nature of the problem directly as a stochastic program. For example, Mulvey (1987) [21] and Mulvey and Vladimirou (1989) [22] [23] formulate asset allocation problems as a stochastic network problem.

The use of stochastic programming techniques has been hampered until recently by the sheer size of practical problems when they are restated as deterministic linear problems. To solve them it was necessary that the number of scenarios representing uncertainties be kept small. Most models developed so far have been single-stage or single-period models, that is to say to the case where the decision making process and the future events (foresight) are restricted to a single time period. Only few attempts have been made to solve practical multi-stage decision making models whose future events are spread over several periods.

Multi-stage planning problems can often be formulated as linear programs with a dynamic matrix structure which, in the deterministic case, appear in a staircase pattern of blocks with non-zero submatrices. These blocks correspond to and are different for different time periods. In the stochastic case, the blocks of coefficients and right hand sides in different time periods are functions of several parameters whose values vary stochastically with dependent and independent distributions which we assume to be known. The resulting problem is a multi-stage stochastic linear program. Even for problems with a small number of stochastic parameters per stage the size of multi-stage problems when expressed in equivalent deterministic form can get so large as to appear intractable. The simplest case and most studied is that with

two stages. Stochastic linear programs were first introduced by Dantzig (1955) [4] and Beale (1955) [1]. Since then it has been studied by many authors, some recent references are Birge (1985) [3], Ermoliev (1983) [10], Frauendorfer (1988) [12], Higle and Sen (1989) [14], Kall (1979) [19], Pereira et al. (1989) [25], Rockafellar and Wets (1989) [28], Ruszczynski(1986) [29], and Wets (1984) [31]. See Ermoliev and Wets (1988) [11] for a survey of different ways proposed to solve the stochastic programs.

A new approach based on Benders decomposition and importance sampling was introduced by Dantzig and Glynn (1990) [5] and developed jointly by them and Infanger (1990) [17]. Our approach turned out to be very powerful. We demonstrated its power by solving several practical large-scale stochastic linear programs with numerous stochastic parameters. Infanger (1991) [18] and Dantzig and Infanger (1991) [7] report on computational results of large-scale problems with up to 52 stochastic parameters, where the deterministic equivalent problem if attempted to express it explicitly would have had several billions of constraints. These problems were two-stage problems or belonged to a restricted class of multi-stage problems which could be reexpressed in the two-stage framework.

2. The Multi-Period Asset Allocation Problem

In this paper we formulate a class of multi-period financial asset allocation problems (Mulvey and Vladimirou (1989) [22]) and show how they can be solved by adaptations of multi-stage stochastic linear programs methodology and software.

At the initial time period 1 a certain amount of wealth is available to a decision maker in assets $i = 1, \ldots, n$ and in cash which we index as asset $n + 1$. We denote $\bar{x}_i, i = 1, \ldots, n+1$ to be the dollar value of the initially available assets. The decision maker has to decide each period how to rearrange his portfolio to achieve best return on his initial investment over time. We consider the problem in discrete time and define time steps $t = 1, \ldots, T$, e.g. by months, with T being the end of the planning horizon.

At each time period t the investor can either hold on to asset i, buy more, or sell off part (or all) of asset i. We denote y_i^t the amount sold of asset i in period t and by

x_i^t the amount of asset i in period t held on to. Selling means decreasing the value x_i^t of asset i and increasing the value of cash, x_{n+1}^t. Also the investor has the choice of using his resulting cash to buy certain amounts of assets i. The amount bought in period t is denoted with z_i^t.

Buying and selling causes transaction costs which we assume to be proportional to the amount of dollar value of asset traded. We denote by $100\nu_i$ the transaction costs (expressed as a percentage) associated with buying one unit of i and with $100\mu_i$ the transaction costs (expressed as a percentage) associated with selling off 1 unit of asset i. Buying 1 unit of asset i requires $1+\nu_i$ units of cash and selling 1 unit of asset i results in $1-\mu_i$ units of cash.

Through buying and selling the investor can restructure his portfolio in each time period t. Once this t-th stage decision is made, the holdings x_i^t, $i=1,\ldots,n+1$ can be calculated. The shares in the portfolio is then kept constant till the next time period. The value of x_i^t is affected by the returns on the market. For example a portfolio x_i^t at time t changes its value to $R_i^t x_i^t$ where R_i^t denotes the return factors from period t to period $t+1$.

At time t, when the decision on rearranging the portfolio has to be made, returns R_i^t, for $i=1,\ldots,n$ are not known to the decision maker with certainty. Only the return on cash, R_{n+1}^t is assumed known. However, we assume we know the probability distributions of R_i^t. The problem is of the "wait-and-see" type. While the decision at t has to be made on the basis of distributions of future returns R_i^t, for $i=1,\ldots,n$, $t=1,\ldots,T$, the values of prior returns R_i^t, $i=1,\ldots,n$, $t=1,\ldots,t-1$ have already been observed. We denote with $R^t=R_i^t$, for $i=1,\ldots,n$ the n-dimensional random vector with outcomes $r^t(\omega_t)$, $\omega_t\in\Omega_t$, with p^{ω_t} the corresponding probability and Ω_t the set of all possible outcomes in t. The random returns R_i^t of period t are mutually dependent and dependent on the random parameters of the previous period.

After the last period T no decision is made. Only the value of the portfolio is determined by adding all values of assets including the last period returns. We call this value v^T. The goal of the decision maker, however, is to maximize $Eu(v^T)$, the expected utility of the value of the portfolio after period T. The utility function

$u(v^T)$ describes the way the investor views risk. If $u(v^T)$ is linear, it describes risk neutrality, if $u(v^T)$ is concave, it models risk averseness. Nonlinear utility functions require non linear programming techniques for the solution of the problem. Our methodology is not restricted to linear problems. However, for the sake of ease and computational speed we approximate the nonlinear function by a piecewise linear function with sufficiently large number of linear segments.

In the model presented here we do not consider shortselling of assets, although this feature could be incorporated easily. We also do not consider borrowing of cash, which also could be incorporated easily. The holdings of assets, as well as the amounts of assets sold or bought have to be positive. In general there are also lower ($\underline{x}$) and upper ($\overline{x}$) bounds on holdings as well as on amounts of assets to be sold ($\underline{y}, \overline{y}$) or to be bought ($\underline{z}, \overline{z}$) which are given by the investor and/or by the market. E.g. a certain asset may only be available up to a certain amount or an investor wants to have a certain asset with at least a certain amount of dollar value in the portfolio. Therefore in general we formulate $\underline{x}_i^t \le x_i^t \le \overline{x}_i^t$, $\underline{y}_i^t \le y_i^t \le \overline{y}_i^t$, $\underline{z}_i^t \le z_i^t \le \overline{z}_i^t$, where $\underline{x}_i^t \ge 0$, $\underline{y}_i^t \ge 0$, $\underline{z}_i^t \ge 0$, x_i^0 given for $i = 1, \ldots, n+1,\ t = 1, \ldots, T$.

We can now state the model:

$t = 1, \ldots, T,\ i = 1, \ldots, n+1,\ r_i^0 x_i^0$ given:

$$
\begin{array}{rcrcrcrcl}
-r_i^{t-1}x_i^{t-1} & + & x_i^t & + & y_i^t & - & z_i^t & = & 0,\ i = 1, \ldots, n \\
-r_{n+1}^{t-1}x_{n+1}^{t-1} & + & x_{n+1}^t & - & \sum_{i=1}^n (1-\mu_i) y_i^t & + & \sum_{i=1}^n (1+\nu_i) z_i^t & = & 0, \\
 & & & & & & & & \\
-\sum_{i=1}^{n+1} r_i^T x_i^T & + & v^T & & & & & = & 0,
\end{array}
$$

$$
\max \quad Eu(v^T)
$$

$$
\underline{x}_i^t \le x_i^t \le \overline{x}_i^t, \quad \underline{y}_i^t \le y_i^t \le \overline{y}_i^t, \quad \underline{z}_i^t \le z_i^t \le \overline{z}_i^t, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T
$$

We describe correlation between asset returns using a factor model. Using factors is common in the financial industry (e.g. Perold (1984) [27]), hence historical data of various factors are commercially available. The idea of the factor model is to relate the vector of asset returns $R^t = (R_1, \ldots, R_n)^t$ to factors $V^t = (V_1, \ldots, V_h)^t$. While the number of assets, n is large, e.g. a model should be able to handle about 500 to

3000 assets, the number of factors h is comparatively small. Factor models used in the financial industry typically involve no more than 20 different time series called factors. The factor matrix $F(n \times h)$ relates R^t to V^t :

$$R^t = FV^t$$

The coefficients of the factor matrix are estimated using regression analyses on historical data. By linear transformations of historical factors the transformed factors can always be determined in such a way that the factors V^t are orthogonal. These factors can then be interpreted as **independent** random parameters assumed normally distributed or log normally distributed. Using the factor model stochastically dependent returns can be generated in the computer by using these stochastically independent factors. We denote the random factor V_i^t by v_i^{t,ω_t}, also denoted as v_i^t, with corresponding probability $p(v_i^t)$, where $p(v_i^t) = \text{prob}\ (V_i^t = v_i^t)$.

We also consider inter-period dependency. For example we may wish to have a higher probability of having a high rate of return in period t if it was high in period $t-1$ than if it was low in period $t-1$. We can model this inter-period dependency as a Markovian type process applied directly on the factors:

$$v_i^t = v_i^{t-1} + \eta_i^t, \quad i = 1, \ldots, h$$

The value of factor i in period t is the sum of the value of factor i in the previous period $t-1$ plus some independent random variation of the factor in t, denoted by η_i^t. The Markovian type model can be estimated based on historical data. Instead of having an additive effect as above we may prefer to have a multiplicative effect by applying the Markovian process directly to the logs of the factors. We haven't explored this alternative.

3. Multi-Stage Stochastic Linear Programs

As one can now see easily, the multi-period asset model proposed fits exactly into the framework of a general class of multi-stage stochastic linear programs with recourse. The factor model for generating dependent returns and the Markovian process for

inter-period dependency define a special class of dependencies between stochastic parameters which we will exploit to solve the problem. Before doing so we state the general problem and the methodology we have developed to solve it.

The multi-stage stochastic linear program can be formulated as follows:

$$
\begin{array}{lllllllll}
min\ z & = & & & & & & & \\
c_1x_1 & + & E(c_2x_2^{\omega_2} & + \cdots + & E(c_{T-1}x_{T-1}^{\omega_{T-1},\ldots,\omega_2} & + & E(c_Tx_T^{\omega_T,\ldots,\omega_2}))\cdots) & & \\
s/t & & & & & & & & \\
A_1x_1 & & & & & & & = & b_1 \\
-B_1^{\omega_2}x_1 & + & A_2x_2^{\omega_2} & & & & & = & b_2^{\omega_2} \\
 & & & \ddots & & & & & \vdots \\
 & & & & -B_{T-1}^{\omega_T}x_{T-1}^{\omega_{T-1},\ldots,\omega_2} & + & A_Tx_T^{\omega_T,\ldots,\omega_2} & = & b_T^{\omega_T} \\
 & & & & & & & & \\
x_1 & , & x_2^{\omega_2} & ,\ldots, & x_{T-1}^{\omega_{T-1},\ldots,\omega_2} & , & x_T^{\omega_T,\ldots,\omega_2} & \geq & 0
\end{array}
$$

$$\omega_t \in \Omega_t, t = 2, \ldots, T$$

The problem is the stochastic extension of a deterministic dynamic linear program. While the first stage parameters c_1, A_1, b_1 are known to the planner with certainty, the parameters of stages $2, \ldots, T$ are assumed known only by their distribution. We assume uncertainty in the coefficients of the transition matrices $B_t^{\omega_t}$, $t = 1, \ldots, T$ and the right hand sides $b_t^{\omega_t}$, $t = 1, \ldots, T$ and assume the coefficients of the technology matrices A_t, $t = 2, \ldots, T$ and the objective function coefficients c_t, $t = 2, \ldots, T$ to be known with certainty. The goal of the planner is to minimize the expected value of present and future costs.

The underlying "wait-and-see" decision making-process is as follows: The decision maker makes a first stage decision $\hat{x}_1$ before observing any outcome of random parameters. Then he waits until an outcome of the second stage random parameters gets realized. The second stage decision then is made based on the knowledge of the realization ω_2 but without observing any outcome of random parameters of stages $2, \ldots, T$, and so forth. As the state (the actual outcome) is carried forward to the following period, the decision tree grows exponentially with the number of stages. We

consider discrete distributions of random parameters with finite number of outcomes, e.g. $\omega_t \in \Omega_t$, $\Omega_t = \{1, \ldots, K_t\}$, $t = 1, \ldots, T$. With K_t being the number of scenarios in period t, the total number of scenarios for all T stages is $\prod_{t=1}^{T} K_t$. The number K_t is expected to be large, as it is computed by the crossing of the sets of possible outcomes of the different random parameters within a period. E.g. the dimension of the random vector in period t is h_t and Ω_t^j contains k_t^j elements; then $K_t = \prod_{j=1}^{h_t} k_j^t$. For example, in the asset allocation problem, consider the case of 20 factors, modeled as random parameters with 5 outcomes each: the number of scenarios per period is $5^{20} \approx 10^{14}$. If there are 3 periods, then the total number of scenarios grows to 10^{28}. The dimensions of an equivalent linear program of an asset allocation problem with a universe of about 500 assets is approximately $5 \cdot 10^{30}$ rows and $1.5 \cdot 10^{31}$ columns. It is of course impossible to write down this linear program explicitly.

It is clear that the multi-period asset allocation problem defined above is a special case of the multi-stage stochastic linear program. The correspondence is as follows: the vector x_t now denotes the vector of all decision variables (holdings, amount to be bought and to be sold) in period t. *Uncertainty* occurs only in the transition matrices B_t which contain in their diagonal the return factors R_i^t. The right hand sides $b_2, \ldots, b_T$ are zero, as well as the objective function coefficients $c_2, \ldots, c_{T-1}$. We now describe the techniques we have developed to solve the multi-stage program.

4. Benders Decomposition

A description of how Benders (1962) [2] Decomposition Algorithm can be applied to solve stochastic linear programs can be found in Van Slyke and Wets (1969) [30], Birge (1985) [3]. Using Benders decomposition we decompose the problem into subproblems of different stages t. In the most general case where there is a dependency of stochastic parameters between stages the number of subproblems is equal to the number of scenarios in each stage t. To distinguish one subproblem from another, each is indexed with $\omega_t, \ldots, \omega_2$, where ω_t is the random event in stage t and $\omega_{t-1}, \ldots, \omega_2$ is the path of previous events which gave rise to the particular subproblems in stage t.

For expository purposes, we assume initially the random events that happen in

one stage are independent of those that happen in the next stage. For example, when the probability of having a high rate of return in period t is the same for all values of rate of return in period $t-1$. In the independent case scenarios $\omega_{t+1} \in \Omega_{t+1}$ in period $t+1$ are identical for each scenario $\omega_t \in \Omega_t$ in period t. The history is only carried forward through optimal decisions $\hat{x}_{t-1}^{\omega_{t-1},\ldots,\omega_2}$ from previous periods. In the special class of Markovian dependency which we described earlier, $B_t^{\omega_{t+1},\omega_t} = B_{t-1}^{\omega_t} + \epsilon_t^{\omega_{t+1}}$, where ϵ_t represents a matrix of random parameters independent of those in period $t-1$.

The idea of using Benders decomposition is to express in each stage t, $t = 1,\ldots,T-1$ and scenario ω_t the expected future costs (the impact of stages $t+1,\ldots,T$) by a scalar θ_t and "cuts", necessary conditions for feasibility and optimality which are expressed only in terms of the stage t decision variables x_t and θ_t. Cuts are initially absent and then sequentially added to the stage t problems. Each scenario subproblem ω_t in stage t collects the information about expected future costs by means of the cuts.

The relation between the stages and scenarios in the decomposed multi-stage problem is summarized as follows:

Stage 1 problem:

$$
\begin{array}{lrcl}
min\ z_1 = & c_1 x_1 + \theta_1 & & \\
s/t & & & \\
\pi_1 : & A_1 x_1 & = & b_1 \\
\rho_1^{l_1} : & -G_1^{l_1} x_1 + \theta_1 & \geq & g_1^{l_1}, \quad l_1 = 1,\ldots L_1 \\
& x_1 \ , \ \theta_1 & \geq & 0
\end{array}
$$

Stage t, $t = 2,\ldots,T-1$, problem:

$$
\begin{array}{lrcl}
min\ z_t^{\omega_t} = & c_t x_t^{\omega_t} + \theta_t^{\omega_t} & & \\
s/t & & & \\
\pi_t^{\omega_t} : & A_t x_t^{\omega_t} & = & b_t^{\omega_t} + B_{t-1}^{\omega_t} \hat{x}_{t-1} \\
\rho_t^{l_t,\omega_t} : & -G_t^{l_t} x_t^{\omega_t} + \theta_t^{\omega_t} & \geq & g_t^{l_t}, \quad l_t = 1,\ldots,L_t \\
& x_t^{\omega_t} \ , \ \theta_t^{\omega_t} & \geq & 0
\end{array}
$$

Stage T problem:

$$\begin{aligned} min\ z_T^{\omega_T} &= c_T x_T^{\omega_T} \\ s/t & \\ \pi_T^{\omega_T}: \quad A_T x_T^{\omega_T} &= b_T^{\omega_T} + B_{T-1}^{\omega_T} \hat{x}_{T-1} \\ x_T^{\omega_T} &\geq 0 \end{aligned}$$

$min\ z_1$ represents the optimal objective function value in the first stage. x_1, θ_1 represent the optimal solution, the vector π_1 denotes the optimal dual prices associated to the original stage 1 constraints, and the scalars $\rho_1^{l_1}$ are the optimal dual prices associated to the cuts, which have been added so far in iterations $l_1 = 1, \ldots, L_1$. The optimal objective function values $min\ z_t^{\omega_t} = min\ z_t^{\omega_t}(\hat{x}_{t-1})$, and the optimal dual prices $\pi_t^{\omega_t} = \pi_t^{\omega_t}(\hat{x}_{t-1})$ associated to the original stage t constraints in stages $t,\ t = 2, \ldots, T$ and the optimal dual prices $\rho_t^{l_t,\omega_t} = \rho_t^{l_t,\omega_t}(\hat{x}_{t-1})$ associated to the cuts in stages $t,\ t = 2, \ldots, T-1$ are all dependent upon $\hat{x}_{t-1}$, the optimal solution passed as input from the previous stages $t-1$. According to the scenario development in the previous stages an optimal solution $\hat{x}_{t-1}$ is actually indexed by the scenario outcomes of all previous stages and is therefore denoted as $\hat{x}_{t-1}^{\omega_{t-1},\ldots,\omega_2}$. For the sake of exposition, we suppress the scenario history and present the optimal solution of subproblems in stage t, scenario ω_t as a function of the input $\hat{x}_{t-1}$.

We compute the expected future costs as $z_{t+1} = E_{\omega_{t+1}} z_{t+1}^{\omega_{t+1}}$, the right hand sides of the cuts as $g_t^{l_t} = E_{\omega_{t+1}}(\pi_{t+1}^{\omega_{t+1}} b_{t+1}^{\omega_{t+1}} + \sum_{l_{t+1}=1}^{L_{t+1}} \rho_{t+1}^{l_{t+1},\omega_{t+1}} g_{t+1}^{l_{t+1}})$ and the coefficients of the cuts as $G_t^{l_t} = E_{\omega_{t+1}} \pi_{t+1}^{\omega_{t+1}} B_t^{\omega_{t+1}}$, where $\rho_T^{\omega_T} = 0$, $G_T^{\omega_T} = 0$, and $g_T^{\omega_T} = 0$.

A subproblem in stage t and in scenario ω_t interacts with its predecessors and descendants by passing forward optimal solutions and backwards cuts. Benders decomposition splits the multi-stage problem into a series of two-stage relations which are overall connected by a nesting scheme. We call the stage t, scenario ω_t problem the current master problem. It receives from its ancestor in period $t-1$ a solution $\hat{x}_{t-1}$. The current scenario is determined by the outcome ω_t of the random parameters in stage t which are reflected in the right hand side $b_t^{\omega_t} + B_{t-1}^{\omega_t} \hat{x}_{t-1}$. As stated above, $\hat{x}_{t-1}$ has a history. The history has to be considered when nesting several stages. Given and subject to $\hat{x}_{t-1}$ we solve the stage t problem in scenario ω_t and

pass the obtained solution $\hat{x}_t^{\omega_t}$ to the descendant problems. By solving all problems $\omega_{t+1} \in \Omega_{t+1}$ (referred to as the universe case) we compute the expected value of the descendant stage costs $z_{t+1} = E_{\omega_{t+1}} z_{t+1}^{\omega_{t+1}}$ and the coefficients $G_t = E_{\omega_{t+1}} \pi_{t+1}^{\omega_{t+1}} B_{t+1}^{\omega_{t+1}}$ and the right hand side $g_t = E_{\omega_{t+1}}(\pi_{t+1}^{\omega_{t+1}} b_{t+1}^{\omega_{t+1}} + \sum_{l_{t+1}=1}^{L_{t+1}} \rho_{t+1}^{l_{t+1},\omega_{t+1}} g_{t+1}^{l_{t+1}})$ of a cut. The cut is added to the current master problem (stage t, scenario ω_t problem) and by solving the problem again another trial solution is obtained.

The optimal solution the of current master problem in stage t, scenario ω_t gives a lower bound, and the expected cost of the trial solution gives an upper bound of the expected costs of all scenarios descendant from the stage t scenario ω_t. If lower bound and upper bound are sufficiently close, the current master problem is said to represent the future expected cost and contains (by means of a sufficient number of cuts) all the information needed from future scenarios. In this case we say the **current master is balanced** with its descendant problems.

Note that the current master problem represents the expected future costs only subject to the trial solution $\hat{x}_{t-1}$ which was passed from its ancestor and subject to the current scenario ω_t. Note also that we have implicitly assumed that the descendant problems in stage $t+1$ are also balanced with their descendant problems in stage $t+2$ by means of having collected a sufficient number of cuts to represent the expected costs of descendant scenarios for $t+2$ on, and so forth. However, note that the solution of the current stage t scenario ω_t problem gives a lower bound of the expected costs of all scenarios descendant from the stage t scenario ω_t problem regardless of having collected a sufficient number of cuts. We shall exploit this fact.

Two properties of cuts are crucial for the solution procedure:

1. In the case of independence of stochastic parameters between stages: The cuts derived from any trial solution $\hat{x}_t^{\omega_t}$ are valid cuts for all subproblems $\omega_t \in \Omega_t$. E.g. the cut: $\theta_t \geq E_{\omega_{t+1}} \pi_{t+1}^{\omega_{t+1}} B_t^{\omega_{t+1}} x_t + E_{\omega_{t+1}} \pi_{t+1}^{\omega_{t+1}} b_t^{\omega_{t+1}}$ is a constraint whose coefficients don't depend on x_t, hence is valid for all values of x_t. To see this, note $\pi_{t+1}^{\omega_{t+1}} = \pi_{t+1}^{\omega_{t+1}}(\hat{x}_t)$ are optimal dual prices that do depend on $\hat{x}_t$ for optimality but they remain **dual feasible** independent of the values of the right hand side as a function of $\hat{x}_t$. The validity of the cuts depends only on the dual feasibility of the

$\pi_{t+1}^{\omega_{t+1}}$. It represents an outer linearization of the future expected cost function $z_{t+1}(x_t)$, evaluated at $\hat{x}_t$. Different scenarios ω_t are in stage t are distinguished by different right hand sides of the original stage t constraints, e.g. $A_t x_t = b_t^{\omega_t} + B_{t-1}^{\omega_t}\hat{x}_{t-1}$. The set of cuts $-G_t^{l_t}x_t + \theta_t \geq g_t^{l_t}, l_t = 1,\ldots,L_t$ represent an outer linearization of the expected future costs independent of scenarios $\omega_t \in \Omega_t$. The outer linearization defined by the set of cuts equals the expected future cost function, if $Ez_{t+1}^{\omega_{t+1}}(\hat{x}_t) = \hat{\theta}_t$, where $\hat{\theta}_t$ is the value of θ_t corresponding to the solution $\hat{x}_t$ of any stage t problem. If $Ez_{t+1}^{\omega_{t+1}}(\hat{x}_t^{\omega_t}) = \theta_t^{\hat{\omega}_t}$, $\omega_t \in \Omega_t$, then a sufficient number of necessary cuts have been generated to represent the expected future costs for all solutions $\hat{x}_t^{\omega_t}$ of scenarios $\omega_t \in \Omega_t$ in stage t and we say stage t is balanced with stage $t+1$.

2. In the case of dependency of stochastic parameters between stages: Cuts now depend on scenario ω_t in period t. Sharing of cuts between different scenario subproblems $\omega_t \in \Omega_t$ is no longer directly possible. However, for additive dependency, (e.g. Markovian type dependency) cuts can be easily adjusted to different scenarios. (See Pereira and Pinto (1989) [26] for additive dependent right hand sides.) For example in the case of the Markovian type dependency which we introduced in the multi-period asset allocation problem $B_t^{\omega_{t+1},\omega_t} = B_{t-1}^{\omega_t} + \epsilon_t^{\omega_{t+1}}$. Here ϵ_t represents a matrix whose elements are functions of random parameters which are independent of the period $t-1$ random parameters. (The elements of ϵ are the independent part of the random returns and are generated by the product $F\,\eta_t$ where η_t is the random change in V_t that generated ϵ_t.) In the case of the additive dependency a cut in stage t and scenario ω_t has the form: $\theta_t \geq [(E_{\omega_{t+1}}\pi_{t+1}^{\omega_{t+1}})B_t^{\omega_t} + E_{\omega_{t+1}}\pi_{t+1}^{\omega_{t+1}}\epsilon_t^{\omega_{t+1}}]x_t + E_{\omega_{t+1}}\pi_{t+1}^{\omega_{t+1}}b_t^{\omega_{t+1}}$. It can be easily seen that the coefficients of the cut consist of a part independent of scenarios ω_t and a dependent part. The cut can be adjusted to different scenarios $\omega_t \in \Omega_t$ by adding the scenario dependent part $(E_{\omega_{t+1}}\pi_{t+1}^{\omega_{t+1}})B_t^{\omega_t}$ according to scenario ω_t. This requires storing of the expected value of the dual variables $E_{\omega_{t+1}}\pi_{t+1}^{\omega_{t+1}}$.

Taking advantage of the above stated properties we actually only need to store one subproblem per stage t. For different scenarios ω_t and different solutions $\hat{x}_{t-1}$ passed from the previous stage we determine the right hand side accordingly. The cuts are valid for all scenarios $\omega_t \in \Omega_t$ in the case of independence of the stochastic

parameters between stages or are adjusted in the gradient according to the actual scenario ω_t in case of Markovian type dependency between stages. Therefore it is easily possible to generate any ω_t subproblem. Future information is represented in the cuts which have been generated so far and can be efficiently used in any scenario $\omega_t \in \Omega_t$ independently from which scenario originated it.

5. Multidimensional Integration

The computation of the expected future costs z_{t+1}, the coefficients G_t and the right hand side g_t of the cuts requires the computation of multiple integrals or multiple sums. The expected value of the second stage costs in period $t+1$ (we suppress the index t for this discussion), e.g. $z = Ez^\omega = E(C)$ is an expectation of functions $C(v^\omega), \omega \in \Omega$, where $C(v^\omega)$ is obtained by solving a linear program. V (in general) is a h-dimensional random vector parameter, e.g. $V = (V_1, \ldots, V_h)$, with outcomes $v^\omega = (v_1, \ldots, v_h)^\omega$. For example V_i represents the value of the i-th factor v_i^ω the observed random outcome. The vector v^ω is also denoted by v, and $p(v^\omega)$ alias $p(v)$ denotes the corresponding probability. Ω is the set of all possible random events and is constructed by crossing the sets of outcomes $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_h$. With P being the probability measure under the assumption of independence the integral $E\ C(V) = \int C(v^\omega) P(d\omega)$ takes the form of a multiple integral $E\ C(V) = \int \cdots \int C(v)p(v)dv_1 \cdots dv_h$, or, in case of discrete distributions, the form of a multiple sum $E\ C(V) = \sum_{v_1} \cdots \sum_{v_h} C(v)p(v)$, where $p(v) = p_1(v_1) \cdots p_h(v_h)$.

The number of terms in the multiple sum computation gets astronomically large and therefore the evaluations of multiple sums by direct summation is not practical. This is especially true because function evaluations are computationally expensive since the evaluation of each term in the multiple sum requires the solution of a linear program. In the following we discuss a scheme for estimating the expected values with a sufficiently low estimation error without having to evaluate each term.

6. Importance Sampling

Monte Carlo Methods are recommended to compute multiple integrals or multiple sums for higher h-dimensional sample spaces (Davis and Rabinowitz (1984) [9] , Glynn and Iglehart (1989) [13]). Suppose $C^\omega = C(v^\omega)$ are independent random variates of v^ω, $\omega = 1,\ldots,n$ with expectation z, where n is the sample size. An unbiased estimator of z with variance $\sigma_{\bar{z}}^2 = \sigma^2/n$, $\sigma^2 = var(C(V))$ is

$$\bar{z} = (1/n)\sum_{\omega=1}^{n} C^\omega.$$

Note that the standard error decreases with $n^{-0.5}$ and the convergence rate of $\bar{z}$ to z is independent of the dimension of the sample space h. We rewrite $z = \sum_{\omega\in\Omega} C(v^\omega)p(v^\omega)$ as

$$\sum_{\omega\in\Omega} \frac{C(v^\omega)p(v^\omega)q(v^\omega)}{q(v^\omega)}$$

by introducing a new probability mass function $q(v^\omega)$ and we obtain a new estimator of z

$$\bar{z} = \frac{1}{n}\sum_{\omega=1}^{n} \frac{C(v^\omega)p(v^\omega)}{q(v^\omega)}$$

by sampling from $q(v^\omega)$. The variance of $\bar{z}$ is given by

$$var(\bar{z}) = \frac{1}{n}\sum_{\omega\in\Omega}\left(\frac{C(v^\omega)p(v^\omega)}{q(v^\omega)} - z\right)^2 q(v^\omega).$$

Choosing $q^*(v^\omega) = \frac{C(v^\omega)p(v^\omega)}{\sum_{\omega\in\Omega} C(v^\omega)p(v^\omega)}$ would lead to $var(\bar{z}) = 0$, which means one could get a perfect estimate of the multiple sum from only one estimation. Practically, however, this is useless since to compute $q(v^\omega)$ we have to know $z = \sum_{\omega\in\Omega} C^\omega p(v^\omega)$, which is what we are trying to compute in the first place.

The result, however, helps to derive a heuristic for choosing q. It should be proportional to the product $C(v^\omega)p(v^\omega)$ and should have a form that can be integrated easily. Thus a function $\Gamma(v^\omega) \approx C(v^\omega)$ is sought, which can be integrated with less effort than $C(v^\omega)$. Additive and multiplicative (in the components of the stochastic vector v) approximation functions and combinations of these are potential candidates for our approximations. Especially for financial investment problems, we have been

getting good results using $C(V) \approx \sum_{i=1}^{h} C_i(V_i)$. We compute q as

$$q(v^\omega) \approx \frac{C(v^\omega)p(v^\omega)}{\sum_{i=1}^{h}\sum_{\omega\in\Omega_i} C_i(v^\omega)}.$$

In this case one has to compute only h 1-dimensional sums instead of 1 h-dimensional sum. The variance reduction depends on how well the approximation function fits the original cost function. If the original cost function has the property of additivity (separability) the multiple sum can be computed exactly by h 1-dimensional sums. If the additive model is a bad approximation of the cost function the only "price" that has to be paid is increasing the sample size. If the observed variance is too high using a starting sample size, the sample size is adjusted higher. Actually we use a variant of the additive approximation function. By introducing $C(\tau)$, the costs of a base case, we make the model more sensitive to the impact of the stochastic parameters v.

$$\Gamma(V) = C(\tau) + \sum_{i=1}^{h} \Gamma_i(V_i), \quad \Gamma_i(V_i) = C(\tau_1, \ldots, \tau_{i-1}, V_i, \tau_{i+1}, \ldots, \tau_h) - C(\tau)$$

We denote this as a *marginal cost model*. τ can be any arbitrary chosen point of the set of values v_i, $i = 1, \ldots, h$. For example we choose τ_i as that outcome of V_i which leads to the lowest costs, ceteris paribus.

Summarizing, the importance sampling scheme has two phases: the preparation phase and the sample phase. In the preparation phase we explore the cost function $C(V)$ at the margins to compute the additive approximation function $\Gamma(V)$. For this process $n_{prep} = 1 + \sum_{i=1}^{h}(k_i - 1)$ subproblems have to be solved. Using $\Gamma(V)$ we compute the approximate importance density

$$q(v^\omega) = \frac{\Gamma(v^\omega)p(v^\omega)}{C(\tau) + \sum_{i=1}^{h}\sum_{\omega\in\Omega_i} \Gamma_i(v^\omega)p(v^\omega)}.$$

Next we sample n scenarios from the importance density and, in the sample phase, solve n linear programs to compute the estimation of $\bar{z}$ using the Monte Carlo estimator. We compute the gradient G and the right hand side g of the cut using the same sample points at hand from the expected cost calculation. See Infanger (1991) [18] for the computation of the cuts and details of the estimation process.

7. The Algorithm

By solving a sample of subproblems ω_{t+1} according to the importance sampling scheme we compute estimates of the expected future costs $z_{t+1}^{\omega_t}$, the gradients $G_t^{l_t}$ and the right hand sides $g_t^{l_t}$ of the cuts in each stage t and scenario ω_t. The objective function value of the solution of each stage t, scenario ω_t subproblem gives a valid lower-bound estimate of the expected costs $z_t^{\omega_t} = c_t\hat{x}_t^{\omega_t} + \hat{\theta}_t^{\omega_t}$ subject to scenario ω_t and subject to $\hat{x}_{t-1}$, the (optimal) solution passed forward from the previous stage. The obtained lower-bound estimate is the tightest lower bound that can be generated, if in stage $t+1$ a sufficient number of cuts have been added to represent the expected future costs with respect to stage $t+1$ for all scenarios $\omega_{t+1} \in \Omega_{t+1}$ and is a weaker lower-bound estimate if there is not a sufficient number of cuts.

We are especially interested in the lower-bound estimate of the first stage costs which we obtain by solving the first stage problem. If the first stage problem is balanced with the second stage, that is, if the cuts added so far to the first stage problem fully represent the expected second stage costs, and if the second stage is balanced with the third stage for all scenarios $\omega_2 \in \Omega_2$ and all values of $\hat{x}_1$, passed to it from the first stage, and so forth till stage $T-1$, then the solution of the first stage problem is the optimum solution of the multi-stage stochastic linear program. In this case the lower bound estimate of z_1 takes on the value of the total expected costs of the multi-stage problem.

To obtain an upper bound of the total expected costs of the multi-stage problem, we evaluate the expected costs of the current first stage trial solution $\hat{x}_1$. This can be accomplished by sampling paths from stages $2,\ldots,T$. For a reference, see Pereira and Pinto (1989) [26]. To efficiently sample a small number of paths to obtain an accurate estimate of the expected costs associated with $\hat{x}_1$, we also use importance sampling. We define a path $\hat{s}^{\omega} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_T)^{\omega}$, $\omega \in \Omega$, where $\Omega = \{\Omega_2 \times \Omega_3 \times \cdots \times \Omega_T\}$, as a sequence of optimal solutions $\hat{x}_t^{\omega_t}$ of stage t scenario ω_t problems, $t = 2,\ldots,T$ and $\hat{x}_1$ being the first stage trial solution. A path is computed by observing the "wait-and-see" requirements: We pass $\hat{x}_1$ to the second stage and solve the second stage problem for scenario ω_2 and obtain the optimal solution $\hat{x}_2^{\omega_2}$. Next we pass the

obtained second stage solution $\hat{x}_2^{\omega_2}$ to the third stage and solve the third stage problem for scenario ω_3 to obtain $\hat{x}_3^{\omega_3}$. We continue in this way until we obtain $\hat{x}_T^{\omega_T}$ in stage T. Note that when solving the stage t problem no future outcomes $\omega_{t+1}, \ldots, \omega_T$ are used. All future information at each stage is solely represented by means of the cuts added in stage t so far. The costs of a path $\hat{s}^\omega$, $C(\hat{s}^\omega)$ is given by $C(\hat{s}^\omega) = \sum_{t=1}^{T} c_t \hat{x}_t^{\omega_t}$. The expected value of the costs of all paths $\hat{s}^\omega$, $\omega \in \Omega$, $E\hat{s}^\omega$ gives an upper bound to the costs of a trial solution $\hat{x}_1$.

We sample paths by applying the importance sampling scheme to the dimensional space of size $\sum_{t=2}^{T} h_t$ of all random parameters $V_{i_t}^t$, $i_t = 1, \ldots, h_t$, $t = 2, \ldots, T$. For sampling paths the importance density $q(V)$ is computed based on the additive marginal approximation function analogous to the way it was defined earlier:

$$\Gamma(V) = C(\tau) + \sum_{t=1}^{T} \sum_{i_t=1}^{h_t} C(\tau_{1,1}, \ldots, \tau_{t,i_t-1}, V_{t,i_t}, \tau_{t,i_t+1}, \ldots, \tau_{T,h_T}) - C(\tau)$$

where $V = (V_1^1, \ldots, V_{h_1}^1, V_1^2, \ldots, V_{h_t}^T)$ and $\tau = (\tau_1^1, \ldots, \tau_{h_1}^1, \tau_1^2, \ldots, \tau_{h_t}^T)$. Sampling paths $\omega \in \Omega$ according to this importance sampling scheme we obtain an equal number of sample points $\omega_t \in \Omega_t$ in stages $t = 2, \ldots, T$. At these sample points we define the current stage t scenario ω_t subproblems and generate cuts to be added at stages $t = 1, \ldots, T-1$ by employing importance sampling as described above for cuts.

The overall procedure works as follows: Solving the stage 1 problem in iteration 1 we obtain a trial solution $\hat{x}_1$ and a lower bound estimate of the expected costs z_1. Now we employ the path sampling procedure to obtain an upper bound estimate of the expected costs z_1. If the upper bound estimate and the lower bound estimate are within a given optimality tolerance, we call the first stage solution the optimal solution of the multi-stage problem, and quit. Otherwise we generate cuts in stages $1, \ldots, T-1$. The path sampling procedure used for the upper bound estimate has produced sample points $\omega_t \in \Omega_t$ in stages $t = 2, \ldots, T$ with corresponding ancestor solutions $\hat{x}_1$ and $\hat{x}_t^{\omega_t}$ in stages $t = 2, \ldots, T-1$ to be passed to the current stage t scenario ω_t problem. Starting at stage $T-1$ and moving backwards till stage 1 we take each sample problem ω_t in stage t and finally the stage 1 problem as the

current master problem and compute cuts by sampling again $\omega_{t+1} \in \Omega_{t+1}$ descendant problems until each scenario problem ω_t in stage t is balanced with stage $t+1$ with regard to ancestor solutions $\hat{x}_{t-1}$ which have been passed from stage $t-1$. Arriving at stage 1 we obtain a new solution $\hat{x}_1$ and a new lower bound estimate. We continue as defined above by sampling new paths for the upper bound estimate. Finally, after a finite number of iterations, upper and lower bound estimates will be sufficiently close. Upper and lower bound estimates can be seen as the sum of *i.i.d.* random terms which for sample sizes of 30 or more can be assumed normally distributed with known (derived from the sampling process) variances. A 95% confidence interval of the obtained solution is computed.

8. Computational Experience

Computational results of using Benders decomposition and importance sampling for two-stage asset allocation problems can be found in Infanger (1991) [18] and Dantzig and Infanger (1991) [7] where we report on the solution of test problems with up to 52 stochastic parameters and a number of universe scenarios of more than 10^{24}. These problems were formulated as two-stage stochastic programs. Using importance sampling and sample sizes between 200 and 600 very accurate results were obtained, e.g. the estimated 95% confidence interval was less than 0.8% on each side based on the optimal objective function value. Additional tests on these examples showed that the ratio of variance reduction obtained by using importance sampling versus crude (naive) Monte Carlo sampling was about 10^{-6}.

Inspired by these results we implemented an earlier version of the methodology described above for the multi-stage case which did not consider dependency between stages. Instead of the path sampling procedure for obtaining upper bounds we implemented a procedure where we sampled points rather than paths which requested the handling of an exponentially expanding decision tree. Therefore even when we used very small sample sizes, the number of stages that was practical to solve was limited.

We did test up to 3-stage problems. FI3 is a 3-stage test problem derived from a 2-stage financial portfolio problem found in Mulvey and Vladimirou (1989) [22]. The

problem is to select a portfolio which maximizes expected returns in future periods taking into account the possibility of revising the portfolio in each period. There are transaction costs and bounds on the holdings and turnovers. Our test-problem covers a planning horizon of 3 periods whereas the original Mulvey-Vladimirou test-problem was a 2-stage problem which compressed all future periods into a single second stage. They solved the stochastic problem by restricting the number of scenarios in Ω.

We assumed the returns of the stocks in the future periods to be independent stochastic parameters with 3 outcomes each. With 13 assets with uncertain returns, the problem had 26 stochastic parameters instead of 39 because after the last stage decision was made, the expected money-value of the portfolio can be evaluated. The number of universe scenarios was $2.5 \cdot 10^{12}$. (The deterministic equivalent formulation of the problem has more than 10^{14} rows and a similar number of columns.) We obtained an estimated optimal solution of the 3-stage stochastic problem using a sample size of only 50 per stage. The optimal objective function value was estimated to be 1.10895 with an estimated 95% confidence interval of 0.004% on the left side and 0.001% on the right side of the obtained objective function value. Thus the optimal objective value lies within $1.10881 \leq z^* \leq 1.10895$ with 95% probability. Note how small the confidence interval is.

9. Conclusion

We have demonstrated how real-world multi-period asset allocation problems can be efficiently solved as multi-stage stochastic linear programs using our approach of combining Benders decomposition and importance sampling. The numerical results obtained so far are very promising: We obtained very accurate solutions for a 3-stage asset allocation test-problem using remarkably small sample sizes.

References

[1] Beale E.M.L. (1955): On Minimizing a Convex Function Subject to Linear Inequalities, *J. Roy. Stat. Soc. 17b, 173-184.*

[2] Benders, J.F. (1962): Partitioning Procedures for Solving Mixed-Variable Programming Problems, *Numerische Mathematic 4, 238-252.*

[3] Birge, J.R. (1985): Decomposition and Partitioning Methods for Multi-Stage Stochastic Linear Programming, *Operations Research 33, 989-1007.*

[4] Dantzig, G.B. (1955): Linear Programming under Uncertainty, *Management Science 1, 197-206.*

[5] Dantzig, G.B. and Glynn, P.W. (1990): Parallel Processors for Planning Under Uncertainty, *Annals of Operations Research 22, 1-21.*

[6] Dantzig, G.B., Glynn, P.W., Avriel, M., Stone, J., Entriken, R., Nakayama, M. (1989): Decomposition Techniques for Multi-Area Generation and Transmission Planning under Uncertainty, EPRI report 2940-1.

[7] Dantzig G.B. and Infanger, G. (1991): Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition, Technical Report SOL 91-4, Department of Operations Research, Stanford University.

[8] Dantzig, G.B. and Madansky M. (1961): On the Solution of Two-Staged Linear Programs under Uncertainty, *Proc 4th Berkeley Symp. on Mathematical Statistics and Probability I, ed. J. Neyman, 165-176.*

[9] Davis, P.J., and Rabinowitz, P. (1984): Methods of Numerical Integration, Academic Press, London.

[10] Ermoliev, Y. (1983): Stochastic Quasi-gradient Methods and Their Applications to Systems Optimization, *Stochastics 9, 1-36.*

[11] Ermoliev,Y. and R.J. Wets Eds. (1988): Numerical Techniques for Stochastic Optimization, Springer Verlag.

[12] Frauendorfer, K. (1988): Solving SLP Recourse Problems with Arbitrary Multivariate Distributions - The Dependent Case, *Mathematics of Operations Research, Vol. 13, No. 3, 377-394.*

[13] Glynn, P.W. and Iglehart, D.L. (1989): Importance Sampling for Stochastics Simulation, *Management Science, Vol 35, 1367-1392.*

[14] Higle, J.L. and Sen, S. (1989): Stochastic Decomposition: An Algorithm for Two Stage Linear Programs with Recourse, to appear in *Mathematics of Operations Research.*

[15] Ho, J.K. and Loute, E. (1981): A Set of Staircase Linear Programming Test Problems, *Math. Programming 20, 245-250.*

[16] Ho, J.K. and Manne, A.S. (1974): Nested Decomposition for Dynamic Models, *Math. Programming. 6, 121-140.*

[17] Infanger, G. (1990): Monte Carlo (Importance) Sampling within a Benders Decomposition Algorithm for Stochastic Linear Programs, Technical Report SOL 89-13R, Department of Operations Research, Stanford University.

[18] Infanger, G. (1991): Monte Carlo (Importance) Sampling within a Benders Decomposition Algorithm for Stochastic Linear Programs, Extended Version: Including Large-Scale Results, Technical Report SOL 91-6, Department of Operations Research, Stanford University, to appear in *Annals of Operations Research.*

[19] Kall, P. (1979): Computational Methods for Two Stage Stochastic Linear Programming Problems, *Z. angew. Math. Phys. 30, 261-271.*

[20] Markowitz H. (1959): Portfolio Selection: Efficient Diversification of Investments, John Wiley and Sons, New York.

[21] Mulvey (1987): Nonlinear Network Models In Finance, *Advances in Mathematical Programming and Financial Planning, JAI Press, 1, 1987, 253.*

[22] Mulvey, J.M. and H. Vladimirou (1989): Stochastic Network Optimization Models for Investment Planning, *Annals of Operations Research 20 (1989) 187-217*

[23] Mulvey, J.M. and H. Vladimirou (1989): Applying the Progressive Hedging Algorithm to Stochastic Generalized Networks, Statistics and Operations Research Series Report, SOR-89-9, Princeton University, Princeton, New Jersey 08544.

[24] Murtagh, B.A. and Saunders, M.A. (1983): MINOS User's Guide, SOL 82-20, Department of Operations Research, Stanford University, Stanford CA 94305.

[25] Pereira, M.V., Pinto, L.M.V.G., Oliveira, G.C. and Cunha, S.H.F. (1989): A Technique for Solving LP-Problems with Stochastic Right-Hand Sides, CEPEL, Centro del Pesquisas de Energia Electria, Rio de Janeiro, Brazil.

[26] Pereira, M.V., Pinto, L.M.V.G. (1989): Stochastic Dual Dynamic Programming, Technical Note, DEE- PUC/RJ - Catholic University of Rio de Janeiro, Caixa Postal 38063 Gávea, Rio de Janeiro RJ CEP 22452 Brazil.

[27] Perold, A. (1984): Large-Scale Portfolio Optimization, *Management Science 30, 10, 1143.*

[28] Rockafellar, R.T. and Wets, R.J. (1989): Scenario and Policy Aggregation in Optimization under Uncertainty, *Mathematics of Operations Research* (to appear).

[29] Ruszczynski, A. (1986): A Regularized Decomposition method for Minimizing a Sum of Polyhedral Functions, *Mathematical Programming 35, 309-333.*

[30] Van Slyke and Wets, R.J. (1969): L-Shaped Linear Programs with Applications to Optimal Control and Stochastic Programming, *SIAM Journal of Applied Mathematics, Vol 17, 638-663.*

[31] Wets R.J. (1984): Programming under Uncertainty: The Equivalent Convex Program, *SIAM J. on Appl. Math. 14, 89-105.*

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1991	Technical Report

4. TITLE AND SUBTITLE

Multi-Stage Stochastic Linear Programs
for Portfolio Optimization

5. FUNDING NUMBERS

N00014-89-J-1659

6. AUTHOR(S)

George B. Dantzig and Gerd Infanger

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Operations Research - SOL
Stanford University
Stanford, CA 94305-4022

8. PERFORMING ORGANIZATION REPORT NUMBER

1111MA

9. SPONSORING MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research - Department of the Navy
800 N. Quincy Street
Arlington, VA 22217

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

SOL 91-11

12a. DISTRIBUTION AVAILABILITY STATEMENT

UNLIMITED

12b. DISTRIBUTION CODE

UL

13. ABSTRACT *(Maximum 200 words)*

The paper demonstrates how multi-period portfolio optimization problems can be efficiently solved as multi-stage stochastic linear programs. A scheme based on a blending of classical Benders decomposition techniques and a special technique, called importance sampling, is used to solve this general class of multi-stage stochastic linear programs. We discuss the case where stochastic parameters are dependent within a period as well as between periods. Initial computational results are presented.

14. SUBJECT TERMS

Portfolio Optimization; Multi-Stage; Linear Programming;
Large-Scale; Stochastic; Importance Sampling

15. NUMBER OF PAGES

21 pp.

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED			SAR